UI REVOLUTIONIZED BY AI AND ITS POTENTIAL BUSINESS APPLICATION

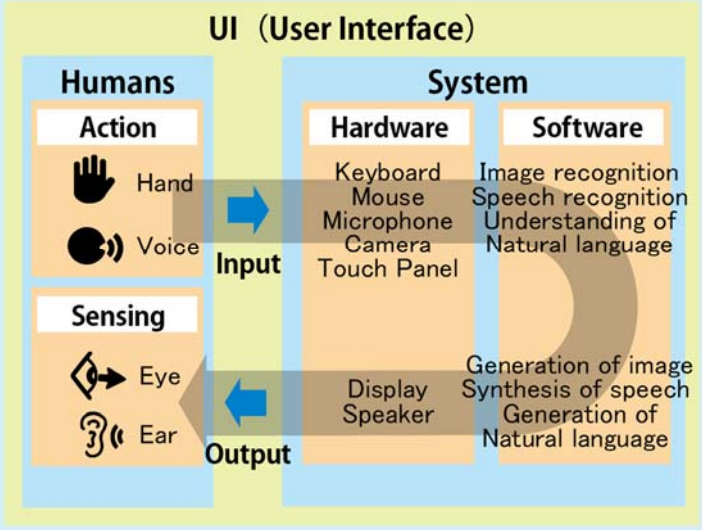
Mitsui & Co. Global Strategic Studies Institute
Technology & Innovation Studies Div.
Digital Innovation Dept.
Koji Naniwada

EMERGENCE OF DEEP LEARNING REVOLUTIONIZES UI

At the 2012 image recognition competition held by the International Society for image recognition, the team led by Professor Geoffrey Hinton of the University of Toronto won overwhelmingly by utilizing deep learning method, far exceeding the recognition accuracy in other teams which utilized the existing machine learning method. Deep learning is the method to make a neural network learn by utilizing a huge amount of learning data in order to generate algorithms automatically. This has triggered improvements in applications of deep learning into various kinds of fields. As a result, functionalities in various kinds of recognition technologies such as image recognition, speech recognition, natural language processing, emotion recognition (image, speech), and action recognition have tremendously improved.

In addition, the combination of these technologies means that various User Interfaces (UIs: the method at the interface where users interact with systems. Refer to Figure 1), such as Chatbot, Voice User Interface (UI), (Augmented Reality/Mixed Reality (AR/MR), are now being put into practical application. These UIs enable interactions with the system to be more natural than the conventional UIs. Those interactions are like communicating with humans. That is why sometimes these UIs are broadly called “Natural User Interfaces” (NUIs). As will be discussed later, Chatbot and Voice UIs are now being put into practical use in various services. Meanwhile, AR/MR and gestures input are expected to be developed into their application phase in the future.

Figure1. Relationship between Humans and UI



Source: Created by MGSSI

Moreover, through utilizing Web API provided by IT vendors, including Google, IBM, and Microsoft, new UIs such as Chatbot and Voice UI have come to be able to be implemented in systems of their own services in a relatively simple way. More specifically, Web API means services provided in cloud computing as well as mechanisms which provide not only AI but also various kinds of functions and services available online by sending simple commands to web servers. By utilizing these functions, system developers can apply new UIs

such as Chatbot and Voice UI to products and services without developing original AI systems in-house from scratch.

NEW UIS AND THEIR TRENDS

(1) Chatbot

Chatbot is a UI which turns communication between humans and systems available into real conversation or interaction with each other in Q&A styles. This function is achieved by analyzing and generating the text with Natural Language Processing (NLP) technology. The Chatbot solution is provided by each IT vendor company. As a typical example of such user support, IBM provides its AI functions through Watson services on its IBM Cloud platform. Conventionally, user support on websites has generally been provided through input via a form, email, and telephone. Meanwhile, with the introduction of Chatbot, users can solve communication problems with the system in real time. Today, this service is widely adopted by websites for e-commerce and financial institutions.

(2) Voice UI

Table 1. Popularity Ranking: Top 20 Rated Amazon Echo Skills
(February 2018, Japanese Market)

Rank	Skill Name	Category	Developer
1	radiko.jp	Music, Audio	radiko
2	Pikachu Talk	Game, Trivia, Accessory	The Pokémon Company
3	LinkJapan eHome	Smart Home	LinkJapan
4	Mameshiba	Novelty, Humor	Dentsu
5	Cookpad	Food, Drink	Cookpad
6	Eki Shiritori powered by Ekispert	Game, Trivia, Accessory	Val Laboratory Corporation
7	Nomura Securities	Business, Finance	Nomura Securities
8	Yahoo! News	News	Yahoo! Japan
9	Yahoo! Weather/Disaster	Weather	Yahoo! Japan
10	Karaoke JOYSOUND	Music, Audio	XING
11	Smart Appliances Controller	Smart Home	RATOC Systems
12	Sleep Support	Life Style	Accela
13	ALC English-Japanese Quiz	Education, Reference	ALC PRESS
14	Sound of Waves	Music, Audio	Excite Japan
15	Find My Phone	Work Efficiency	TrackR
16	528 Ultimate Healing Music	Music, Audio	XING
17	JR-EAST Transport Information	Travel, Transportation	JR-EAST
18	Support for Good Sleep	Life Style	Accela
19	Aruarusan	Game, Trivia, Accessory	YOSHIMOTO ROBOT LABORATORY
20	Ekispert for Web	Travel, Transportation	Val Laboratory Corporation

Source: Created by MGSSI from data based on the announcement by Amazon

Voice UI functions by converting voice signals acquired through a microphone to text through speech recognition, analyzing the meaning of that text by utilizing natural language processing, and then, responding to that content through a speaker. The typical example of this is the voice assistant technology “Alexa” mounted on the AI speaker, “Amazon Echo,” which was released in November 2014, and is now showing rapid prevalence. With the utilization of voice UI, Alexa enables the execution of various tasks such as playing music, management of schedules, shopping, provision of news and weather information. Amazon Echo is an open platform, and therefore each company can develop applications freely for the purpose of utilizing Echo. These applications are called “Skills”, and each company releases various skills (Table 1). Moreover, Amazon is aggressively licensing Alexa to other manufacturers. The biggest advancement in adopting Alexa can be seen in the automobile sector. Automobile manufacturers such as Volkswagen and Ford have announced the adoption of Alexa as their voice UI. In the future, automobiles equipped with “Alexa” are planned to be launched on the market.

(3) AR/MR

AR is the UI which displays the information combined with the vision of the real world. Meanwhile, MR connects the vision of the real world with that of the cyber world. This is the area which attracts much attention today amidst the growing importance of the virtual world. AR/MR needs high-speed image recognition and processing of high resolution images. Therefore, under the current status, it is difficult to provide those services by Web API due to the constraint on communications in terms of bandwidth, delays, and costs. In that sense, Google and Apple have released those services in the form of a Software Development Kit (SDK). SDK means the development environment which makes the development of applications easier. By utilizing this, system developers can develop applications with AR technology relatively quickly. This AR technology is utilizing the cameras of smart phones. If you hold your smartphone over the visible real world in front of you, you can draw the virtual world with which the real world is overlapped. With this function, you can examine the scale and design by locating arbitrary furniture (Fig. 2). If you overlap the guide information with the specified position inside the store with image recognition, this application can guide you, showing the route to the desired product inside the store and the detailed place (Figure 3). In the future, such services are expected to spring up from each company one after another.

Figure 2. Smartphone App “Housecraft” with SDK Released by Apple

Source: Apple Website

Figure 3. Store Guide System with SDK Released by Google

Source: Google Official Twitter Account

(4) Gesture Input

Gesture Input is the UI which enables the input into the system by detecting motions of users' hands and eyes with the technology of image recognition and action recognition. This UI is attracting attention because it can be used for situations where voice output is difficult, in particular, in the case of operating TVs and smartwatches. Samsung is now selling a TV which enables operations such as changing channels and the sound volume with specific hand gestures. Moreover, Google is trying to adopt this function, as seen with the adding of the gesture input function to the OS "Android Wear" for smart watches. This function is activated by the wearer moving his or her arm.

NEW UI'S IMPACT ON BUSINESS

The emergence of these new UIs is likely to have a huge impact on various kinds of businesses. In fact, the smartphone, which uses a UI for touch panel operation, has already appeared in the market. This has led to the availability of operating and browsing maps and photos on a large screen comfortably. Accordingly, smartphone applications have greatly expanded, compared with the existing mobile phones. That means more intuitive operation has become available. Because of this, children and seniors also have come to be able to use such smartphones, and as a result the number of people targeted for such services is increasing. Companies such as Uber, LINE, Mercari, which optimized their own services with these UIs ahead of others, have drastically boosted their business in line with the prevalence of smartphones. When considering such trends, the optimization of services in accordance with the change and evolution in UIs is recognized as having a disruptive impact on the existing business players. For example, with the prevalence of voice UIs, the ordering process for e-commerce purchases will possibly be implemented mainly in living rooms, although ordering on smartphones is now mainstream. Amazon is considered to continue to be a winner in the e-commerce market because it is ahead of other companies in the category of voice UIs. Meanwhile, Google is now increasing the market share of the AI speaker, and thereby has a high possibility to become a major player in the e-commerce market following Amazon by leveraging the production of AI speakers. In addition, the emergence of voice UIs will drastically change the mechanism of infotainment (a combination of information and entertainment) in vehicles. Conventionally, the business of infotainment has been led by automobile manufacturers. However, there is a likelihood that Amazon and Google will lead such business. In other words, it is highly likely that the

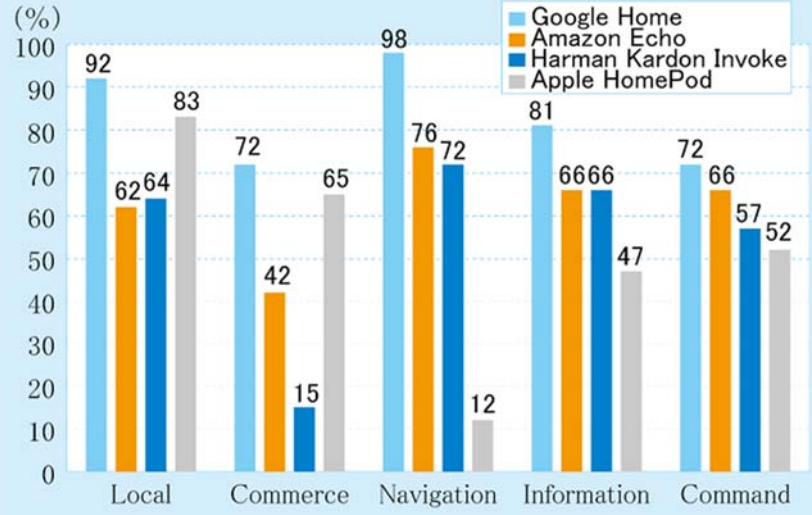
emergence of new UIs such as voice UIs and AR/MR will change the existing framework of businesses, which will result in contributing to the creation of new business as well as the emergence of business players.

Moreover, utilization of these UIs will enable the collection of data such as the conversation (text, voice) data and interior image data, which have hitherto been difficult to obtain. The following are the leading cases of Amazon and Google.

Amazon is far ahead of other companies in the category of voice UI. In specific terms, thanks to production of the Echo series, Amazon is promoting the voice data of living rooms. In the conventional e-commerce of Amazon, only one account has been used for purchase transactions per family in many cases. Therefore, it was difficult to make a distinction in the data for each family member. Since October 2017, the US version of Amazon Echo has corresponded to the function of speaker recognition, and therefore acquisition of individual data has become available. Because of this, more detailed marketing and targeted advertisement becomes available. Recently, Amazon has been focusing on the advertising business. According to eMarketer, a market research company, sales related to advertising in 2017 will be expected to reach 1.65 billion USD, 48% up from the previous year.

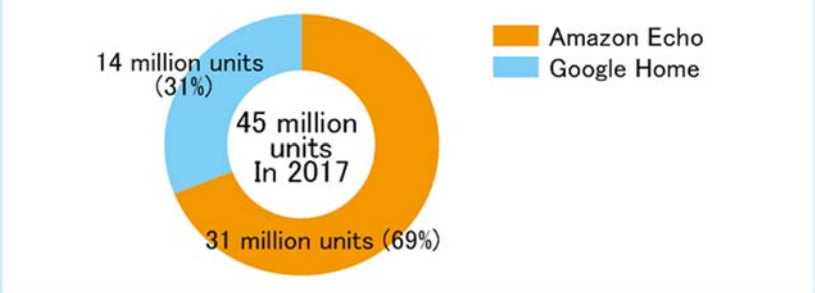
Google is equipped with a strong platform: Android which is an OS for smartphones, and is trying to catch up even in the development of voice UI in which Amazon has taken the lead. According to the benchmark indicated by the research company, the AI speaker “Google Home” released in the summer of 2017 outclasses Amazon in many functions (Figure 4). The market share of AI speaker as of 2017 amounts to 15% in the whole world, and 31% in the U.S. (Figure 5). Also, Google puts much effort into launching services by utilizing AR technology, and is promoting to prepare interior image data for providing services such as indoor navigation and product guides in store. If AR users increase, a huge amount of interior and product images can be expected to be collected in the future.

Figure 4. Benchmark Results of AI Speaker



[Note]
Local: Area information, Commerce: Shopping, Navigation: Root search, Information: Response for general inquiries about share prices, etc., Command: Instructions for alarm setting, etc.
Source: Created by MGSSI based on data from Loup Ventures (February 2018)

Figure 5. Share of AI Speaker

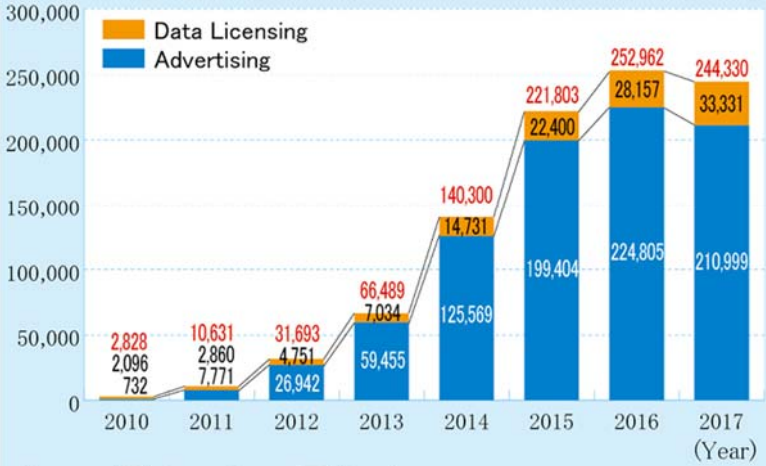


Source: Created by MGSSI based on the CIRP research information

FUTURE OUTLOOK

The evolution of UIs with the utilization of AI is expected to continue. In particular, the utilization of Web API and SDK has come to make UIs implemented in a relatively simple way. That is to say, utilizing Web API and SDK can realize more intuitive and user friendly new services, which have conventionally been difficult to achieve. This will possibly lead to great change in the forms of products and services, while players in businesses will greatly change. Furthermore, with the rise of Voice UI, input with hands and confirmation with eyes has become useless, and therefore, there is a high likelihood that new products and services utilizing Voice UI in market categories such as automobiles, bicycles, and sports will emerge, and new markets will be established. Also, as another effect, data such as conversational (text, voice) data and indoor image data, which were difficult to obtain in the past, can now be collected. Thanks to the prevalence of smartphones since around 2009, much data has come to be collected. Here, we can see the example of monetizing collected data with the company's own service. More specifically, NTT DOCOMO INC, has started to provide the service of forecasting demand for taxi operators since 2017 based on the position information acquired from its own users. Meanwhile, Twitter, Inc. offers a search service for its own SNS data for a charge. In that sense, this data licensing business has become the second largest source of revenue next to the advertisement business (Figure 6). In 2017, advertising revenue has declined, while data licensing business has resulted in strong earnings and Twitter, Inc. has achieved a surplus for the first time since its founding. In the future, it is likely that such business cases as launches of new services by companies to monetize data collected by themselves will increase even for services which utilize new UIs.

Figure 6. Shift in Sales of Twitter, Inc.



Source: IR information of Twitter, Inc.

However, in many cases, it requires a long period of time from the launching of attractive services and enclosing users, to monetizing them by utilizing the data acquired from such users. For example, in the case of NTT DOCOMO INC, it took about 8 years from the full-fledged prevalence of smartphones to commercialization of a demand forecasting service for taxi operators. Even in the case of Twitter, Inc., it required about 7 years from launching the full-fledged services in 2010 to achieving a profit. While maintaining a long-term viewpoint, the immediate start of services optimized for new UIs and enclosure of users is required. In the future, this enhances the potential to create new businesses with the utilization of user data. On top of that, with greater attention being paid to privacy-related matters in recent years, and various governmental authorities strengthening regulations related to privacy, such as the General Data Protection Regulation (GDPR), we need to consider new business creation.